

Abstract

A method of determining data placement for a distributed storage system with a selection of a heuristic class for placing data objects onto nodes of the distributed storage system. The heuristic class meets a performance requirement and provides a replication cost that is within an allowable limit of a minimum replication cost. The method continues with an instantiation of a data placement heuristic selected from a range of data placement heuristics according to the heuristic class. According to one embodiment, the method concludes with the instantiation of the data placement heuristic. According to another embodiment, the method concludes with an evaluation of a placement of the data objects onto the nodes made according to the data placement heuristic. According to another embodiment, the method continues by iteratively performing the selection of the heuristic class, the instantiation of the data placement heuristic, and the evaluation of the data placement heuristic.